UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/954,979	09/17/2001	Jon Rong-Wei Yi	01997-294001	7698

26161 7590 05/19/2004

FISH & RICHARDSON PC
225 FRANKLIN ST
BOSTON, MA 02110

EXAMINER	
VO, HUYEN X	

ART UNIT	PAPER NUMBER
2655	6

DATE MAILED: 05/19/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No. 09/954,979	Applicant(s) YI ET AL.	
	Examiner Huyen Vo	Art Unit 2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *9/17/2001*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-18* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-18* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *17 September 2001* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *5*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 1-13 and 18 are rejected under 35 U.S.C. 102(b) as being anticipated by Kuhn et al. (US Patent No. 6029132).

1. Regarding claims 1 and 18, Kuhn et al. disclose a method for selecting segments from a corpus of source utterances for synthesizing a target utterance (figure 1) and a software stored on a computer-readable medium for causing a computer to perform functions comprising selecting segments from a corpus of source utterances for synthesizing a target utterance (*the operation of figure 1 can be implemented in software*), wherein selecting the segments comprising:

searching a graph in which each path through the graph identifies a sequence of segments of the source utterances and a corresponding sequence of unit labels that characterizes a pronunciation of a concatenation of that sequence of segments, each path being associated with a numerical score that characterizes a quality of the sequence of segment (*referring to figures 1-3, Text-Based Pronunciation Generator 16 searches the Decision Tree 10 to generate a list of pronunciation 18 representing*

possible pronunciation candidates. The Phoneme-Mixed Tree Score Estimator 20 searches the Phoneme-Mixed Decision Three 12 to access the viability of each pronunciation in list 18);

wherein searching the graph includes matching a pronunciation of the target utterance to paths through the graph, and selecting segments for synthesizing the target utterance based on numerical scores of matching paths through the graph (*col. 5, ln. 32-60*).

2. Regarding claim 2, Kuhn et al. further disclose that selecting segments for synthesizing the target utterance includes identifying a path through the graph that matches the pronunciation of the target utterance and selecting the sequence of segments that is identified by the determined path (*col. 2, ln. 66 to col. 3, ln. 6 and col. 5, ln. 33-60*).

3. Regarding claims 3 and 4, Kuhn et al. further disclose that determining the path includes determining a best scoring path through the graph (*col. 4, ln. 30-38*) and determining the best scoring path involves using a dynamic programming algorithm (*col. 4, ln. 30-38*).

4. Regarding claim 5, Kuhn et al. further disclose a method for concatenating the selected sequence of segments to form a waveform representation of the target utterance (*col. 6, ln. 1-8*).

5. Regarding claim 6, Kuhn et al. further disclose that selecting the segments for synthesizing the target utterance includes determining a plurality of paths through the graph that each matches the representation of the pronunciation of the target utterance (*tables 18 and 22, each pronunciation represents a particular path*).

6. Regarding claim 7, Kuhn et al. further disclose that selecting the segments farther includes forming a plurality of sequences of segments, each associated with a different one of the plurality of paths (*the result recorded in tables 18 and 22, with each pronunciation represents a particular path*).

7. Regarding claim 8, Kuhn et al. further disclose that selecting the segments further includes selecting one of the sequences of segments based on characteristics of those sequences of segments not determined by the corresponding sequences of unit labels associated with those sequences (*col. 5, ln. 33-60*).

8. Regarding claim 9, Kuhn et al. further disclose a method for forming a representation of a plurality of pronunciations of the target utterance (*tables 18 and 22 in figure 1*), and wherein searching the graph includes matching any of the pronunciations of the target utterance to paths through the graph (*the operation of elements 10, 16 and elements 20, 50 in figure 1*).

9. Regarding claim 10, Kuhn et al. further disclose a method for forming a representation of the pronunciation of the target utterance in terms of alternating unit labels and transitions labels (*col. 4, ln. 30-38*).

10. Regarding claim 11, Kuhn et al. further disclose that the graph includes a first part that encodes a sequence of segments and a corresponding sequence of unit labels for each of the source utterances (col. 3, ln. 18-25), and a second part that encodes allowable transitions between segments of different source utterances and encodes a transition score for each of those transitions (col. 4, ln. 30-38); and

matching the pronunciation of the target utterance to paths through the graph includes considering paths in which each transition between segments of different source utterances identified by that path corresponds to a different subpath of that path that passes through the second part of the graph (*col. 4, ln. 30-38, different possible combinations of pronunciations are constructed*).

11. Regarding claim 12, Kuhn et al. further disclose selecting the segments for synthesis includes evaluating a score for each of the considered paths that is based on the transition scores associated with the subpaths through the second part of the graph (*col. 4, ln. 30-38, evaluating by determining and selecting the n-best candidates*).

12. Regarding claim 13, Kuhn et al. further disclose that a size of the second part of the graph is substantially independent of a size of the source corpus, and a complexity

of matching the pronunciation through the graph grows less than linearly with the size of the corpus (*node leaf in figures 2-3 represents the second part of the graph. Comparing the size of the node leaf with the whole corpuses 10 and 10, the size of the node leaf is substantially smaller than the corpuses. And thus, all possible combinations of pronunciations at a particular node leaf are much less than the size of the corpuses*).

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 14-17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kuhn et al. (US Patent No. 6029132) in view of Mohri et al. (US Patent No. 6243679).

13. Regarding claim 14, Kuhn et al. further disclose a method of claim 1 further comprising: providing the corpus of source utterances (col. 6, ln. 1-7), forming the graph, including forming a first part of the graph that encodes a sequence of segments and a corresponding sequence of unit labels for each of the source utterances (*col. 3, ln. 18-25*), and forming a second part that encodes allowable transitions between segments of different source utterances and encodes a transition score for each of those transitions (*col. 4, ln. 30-38, all possible combinations of pronunciations*).

Kuhn et al. do not disclose that each source utterance being segmented into a sequence of segments, each consecutive pair of segments in a source utterance forming a segment boundary, and each speech segment being associated with a unit label and each segment boundary being associated with a transition label.

However, Mohri et al. teach that each source utterance being segmented into a sequence of segments, each consecutive pair of segments in a source utterance forming a segment boundary (col. 4, ln. 1-31), and each speech segment being associated with a unit label and each segment boundary being associated with a transition label (col. 4, ln. 1-31). The advantage of using the teaching of Mohri et al. in Kuhn et al. is to interpret the input sequence more correctly.

Since Kuhn et al. and Mohri et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Kuhn et al. by incorporating the teaching of Mohri et al. in order to interpret the input sequence more correctly.

14. Regarding claim 15, Kuhn et al. further disclose that forming the second part of the graph is performed independently of the utterances in the corpus of source utterances (*referring to figures 2-3, all possible combinations of pronunciations is constructed with phonemes at that particular node leaf. Thus, second part of the graph is independent of the corpus*).

15. Regarding claim 16, Kuhn et al. further disclose a method of adding pronunciations to the corpus through the training phase (col. 6, ln. 1-7), but do not specifically disclose a method for augmenting the corpus of source utterances with additional utterances; and augmenting the graph including augmenting the first part of the graph to encode the additional utterances, and linking the augmented first part to the second part without modifying the second part based on the additional utterances. However, it would have been obvious to one of ordinary skill in the art at the time of invention to readily recognize that one can add new pronunciations to the corpus through the training process. The new pronunciations would then be used to represent the pronunciations of the input sequence and all possible combinations of pronunciations at a particular node leaf. This enables the system to personalize the synthetic speech.

16. Regarding claim 17, Kuhn et al. do not disclose that the graph is associated with a finite-state transducer which accepts input symbols that include unit labels and transition labels, and that produces identifiers of segments of the source utterances, and wherein searching the graph is equivalent to composing a finite-state transducer representation of a pronunciation of the target utterance with the finite-state transducer with which the graph is associated.

However, Mohri et al. teach that the graph is associated with a finite-state transducer which accepts input symbols that include unit labels and transition labels, and that produces identifiers of segments of the source utterances (col. 10, ln. 28 to col.

11, ln. 67), and wherein searching the graph is equivalent to composing a finite-state transducer representation of a pronunciation of the target utterance with the finite-state transducer with which the graph is associated (col. 11, ln. 31-67). The advantage of using the teaching of Mohri et al. in Kuhn et al. is to achieve time and space minimization efficiencies.

Since Kuhn et al. and Mohri et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art at the time of invention to modify Kuhn et al. by incorporating the teaching of Mohri et al. in order to achieve time and space minimization efficiencies (col. 1, ln. 60 to col. 2, ln. 2).

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Huyen Vo whose telephone number is 703-305-8665. The examiner can normally be reached on M-F, 9-5:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached on 703-305-4827. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Examiner Huyen X. Vo

May 5, 2004

***

DORIS H. TO
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600